Name:_____ Class:_____ Date:_____

Tutorial Ch 2

Indicate the answer choice that best completes the statement or answers the question.

1. A researcher wishes to determine whether the rate of water flow (in liters per second) over an experimental soil bed can be used to predict the amount of soil washed away (in kilograms). The researcher measures the amount of soil washed away for various flow rates and, from these data, calculates the least-squares regression line to be

Amount of eroded soil = 0.4 + 1.3 × (flow rate)

One of the flow rates used by the researcher was 0.3 liter per second, and for this flow rate the amount of eroded soil was 0.8 kilogram. These values were used in the calculation of the least-squares regression line. What is the residual corresponding to these values?

a. 0.01
b. –0.01
c. 0.5
d. –0.5

2. Are avid readers more likely to wear glasses than those who read less frequently? Three-hundred men in Ohio were selected at random and characterized as to whether they wore glasses and whether the amount of reading they did was above average, average, or below average. The results are presented in the following table.

Amount of reading	Glasses? Yes	Glasses? No
Above average	47	26
Average	48	78
Below average	31	70
Total	126	174

What is the proportion of men in the sample who wear glasses?

a. 0.24
b. 0.37
c. 0.42
d. 0.64

3. Data were obtained from the A&W website about the total fat (in grams) and the protein content (in grams) for various items on the menu. Some summary statistics are provided.

Item	**Total fat (grams)**	**Protein (grams)**
Kid's Cheeseburger	24	23
Kid's Hamburger	22	21
Original Bacon Cheeseburger	33	27
Original Bacon Double Cheeseburger	48	45
Original Double Cheeseburger	42	40
Papa Burger	42	41

	Total fat (grams)	**Protein (grams)**
Mean	35.167	32.833
Standard Deviation	10.591	10.362

Correlation $r = 0.983$

The slope of the least-squares regression line for total fat on protein is _____.

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

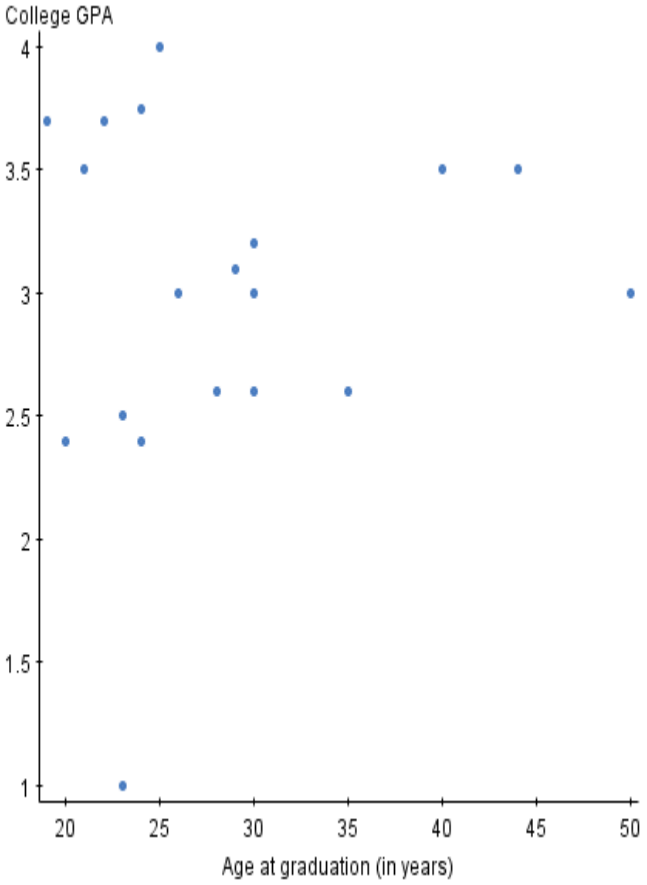
Tutorial Ch 2

a. –0.998
b. 1.005
c. 0.962
d. 2.170
e. 0.966

4. In order for us to examine the relationship between two variables, the variables must be measured from the same _____.

a. cases
b. labels
c. units
d. values

5. The scatterplot below displays data collected from 20 adults on their age and overall GPA at graduation.



There appear to be outliers in the data set.

a. True
b. False

6. Data were obtained from the A&W website about the total fat (in grams) and the protein content (in grams) for various items on its menu. Some summary statistics are provided.

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

Tutorial Ch 2

Item	Total fat (grams)	Protein (grams)
Kid's Cheeseburger	24	23
Kid's Hamburger	22	21
Original Bacon Cheeseburger	33	27
Original Bacon Double Cheeseburger	48	45
Original Double Cheeseburger	42	40
Papa Burger	42	41

	Total fat (grams)	Protein (grams)
Mean	35.167	32.833
Standard Deviation	10.591	10.362

Correlation $r = 0.983$

The intercept for the least-squares regression line of total fat on protein is _____.

a. –0.998
b. 1.005
c. 0.962
d. 2.170
e. 0.966

7. If you have two categorical variables, one way to study them is to use a _____.

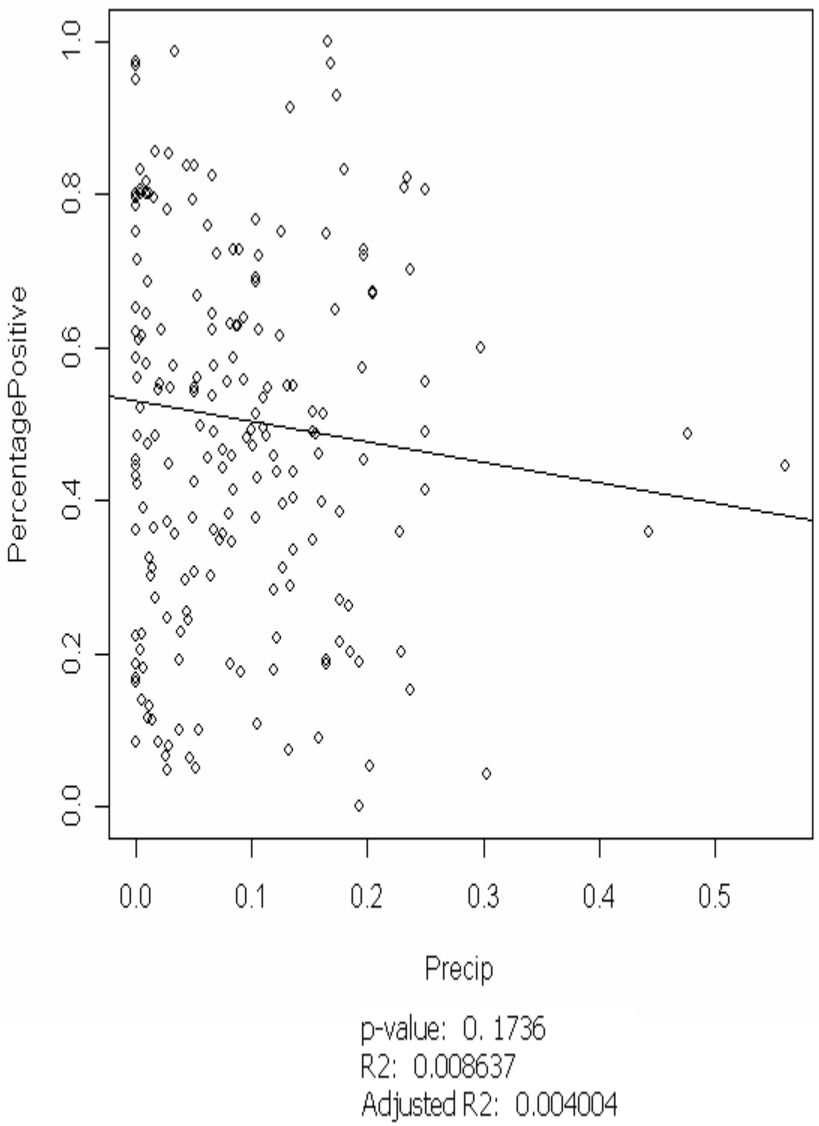
a. scatterplot
b. regression line
d. two-way table

8. Malaria is a leading cause of infectious disease and death worldwide. It is also a familiar example of a vector-borne disease that could be greatly affected by the influence of climate change. The scatterplot shows total precipitation (in mm) in select cities in West Africa on the *x* axis, and the percent of people who tested positive for malaria in the select cities on the *y* axis, in the year 2000.

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

Tutorial Ch 2



There is a strong linear relationship between percent of people who tested positive for malaria and precipitation.

a. True

b. False

9. The value of r^2 ranges between:

a. –1 and 1.

b. 0 and 1.

c. 0 and 100.

d. –100 and 100.

10. A review of voter registration records in a small town yielded the following data for the numbers of males and females registered as Democrat, Republican, or some other affiliation.

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

Tutorial Ch 2

Affiliation	Male	Female
Democrat	300	600
Republican	500	300
Other	200	100

What proportion of registered Democrats are male?

a. 0.15
b. 0.30
c. 0.33
d. 300

11. Correlations based on averages tend to be _____ correlations based on individuals.

a. higher than
b. lower than
c. the same as

12. A review of voter registration records in a small town yielded the following data for the numbers of males and females registered as Democrat, Republican, or some other affiliation.

Affiliation	Male	Female
Democrat	300	600
Republican	500	300
Other	200	100

What proportion of all voters are male and registered as a Democrat?

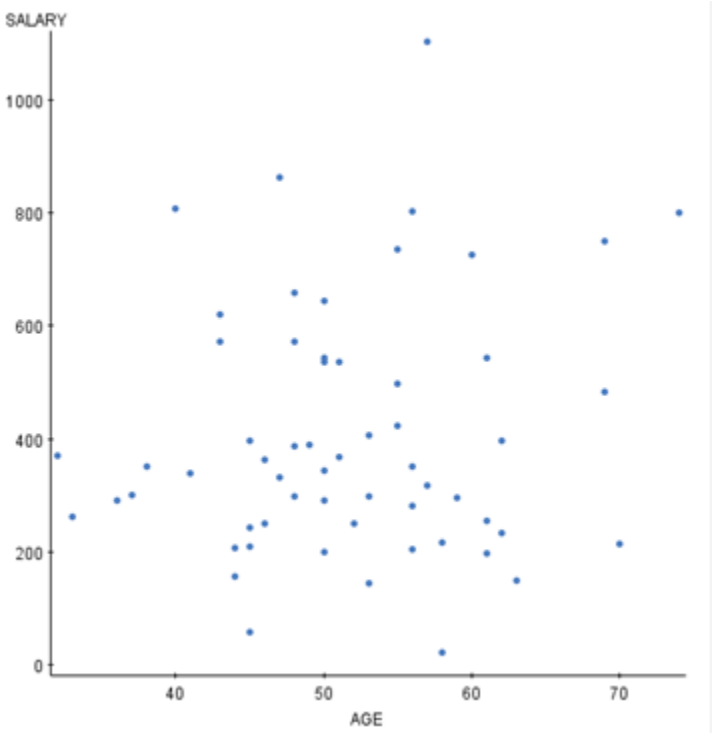
a. 0.15
b. 0.30
c. 0.33
d. 300

13. Is age a good predictor of salary for a CEO? Sixty CEOs between the ages of 32 and 74 were asked their salary (in thousands). The results of a statistical analysis are shown below.

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

Tutorial Ch 2



Simple linear regression results:
Dependent Variable: SALARY
Independent Variable: AGE
SALARY = 242.70212 + 3.1327114 AGE
Sample size: 59
R (correlation coefficient) = 0.1276
R-sq = 0.016270384
Estimate of error standard deviation: 220.64246

Analysis of variance table for regression model:

Source	DF	SS	MS	F-stat	*P*-value
Model	1	45896.027	45896.027	0.9427509	0.3357
Error	57	2774936.2	48683.094		
Total	58	2820832.2			

Is age a good predictor of salary?

a. Yes, the intercept is high.

b. Yes, the correlation is high.

c. No, the intercept is too low.

d. No, the correlation and r^2 are low.

14. Colorectal cancer (CRC) is the third most commonly diagnosed cancer among Americans (with nearly 147,000 new cases annually) and the third leading cause of cancer death (with over 50,000 deaths annually). Research was done to determine whether there is a link between obesity rate and CRC mortality rate among African Americans in the United States by county. Below are the results of a least-squares regression analysis from the software *StatCrunch*.

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

Tutorial Ch 2

Simple linear regression results:
Dependent Variable: Mortality.rate
Independent Variable: Obesity.rate
Mortality.rate = 13.458199 – 0.21749489 Obesity.rate
Sample size: 3098
R (correlation coefficient) = –0.0067
R-sq = 4.5304943E-5
Estimate of error standard deviation: 111.20661

Parameter estimates:

Parameter	Estimate	Std. Err.	Alternative	DF	T-Stat	*P*-Value
Intercept	13.458199	15.9797735	≠ 0	3096	0.84220207	0.3997
Slope	−0.217494 9	0.5807189	≠ 0	3096	–0.37452698	0.708

Analysis of variance table for regression model:

Source	DF	SS	MS	F-stat	*P*-value
Model	1	1734.7122	1734.7122	0.14027046	0.708
Error	3096	3.8287952E7	12366.91		
Total	3097	3.8289688E7			

The correlation between obesity rate and CRC mortality rate is _____.

a. very strong
b. very weak
c. moderately strong
d. moderately weak

15. 130. A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $10,000 or $20,000 per year. The numbers of male and female managers at each salary level and the numbers of male and female workers at each salary level are given in the table below.

Income	Male Managers	Female Managers
$100,000	80	20
$200,000	20	30
Income	Male Works	Female Workers
$10,000	30	20
$20,000	20	80

What is the proportion of male managers who make $200,000 per year?

a. 0.067
b. 0.133
c. 0.2
d. 0.4

16. Which set of two variables is most likely to have a cause-and-effect relationship?

a. The height of a person and the weight of a person

Copyright Macmillan Learning. Powered by Cognero.

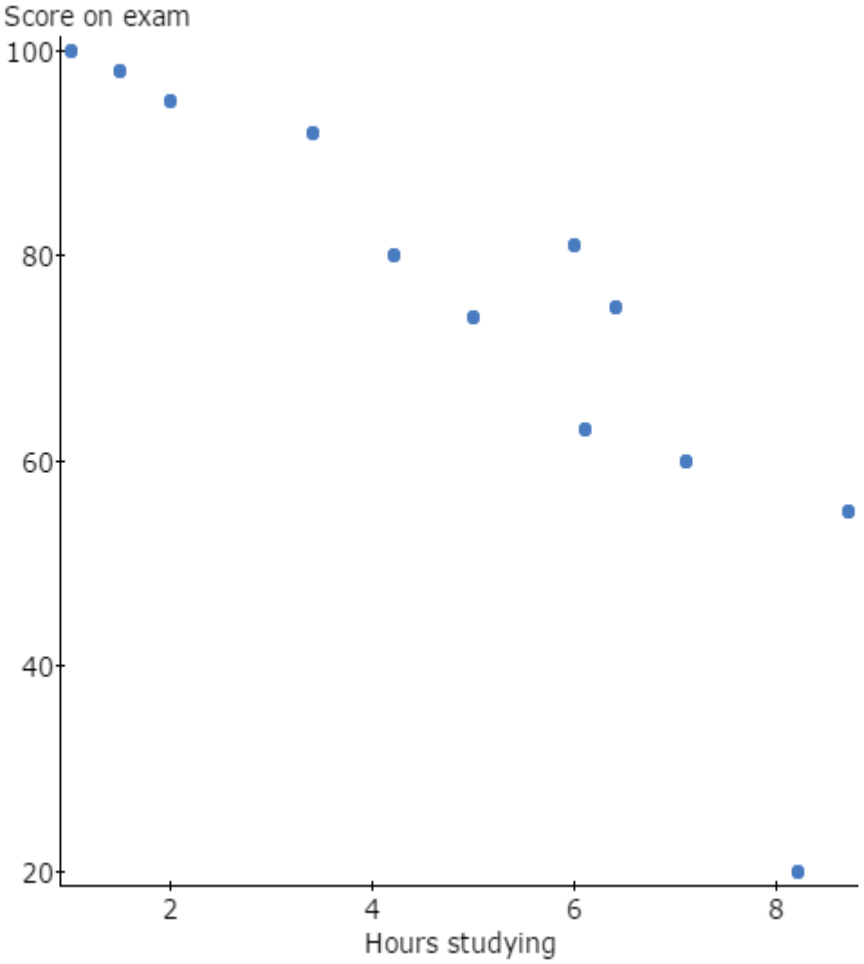
Tutorial Ch 2

b. The weight of a box and the postage rate one has to pay to ship the box to California

c. The make of a car and the mileage of the car

d. The age of a teacher and the income of the teacher

17. The scatterplot illustrates data from a basic statistics class. Students in the class were asked to provide the amount of time (in hours) they spent studying for the first exam. The professor then made a scatterplot to present the relationship between the number of hours a student studied and the score (from 0–100, with 100 being the best score) that the student received on the first exam. How would you interpret this scatterplot?



a. Students who studied the least amount of time received the highest grades. Therefore, they should not study long on a statistics exam if they want to receive a high grade.

b. Students who studied the most received the highest grades. Therefore, they should study several hours to receive the highest exam scores.

c. The correlation is likely a nonsense correlation caused by a lurking variable. Students who received higher scores probably did not need to study as much because they were doing better in the course than students who received lower scores.

d. None of the above

18. An experiment is conducted to study the bonding strength of adhesives that contain varying amounts of a particular chemical additive. Wafers of a specified material are glued together using the adhesive with each

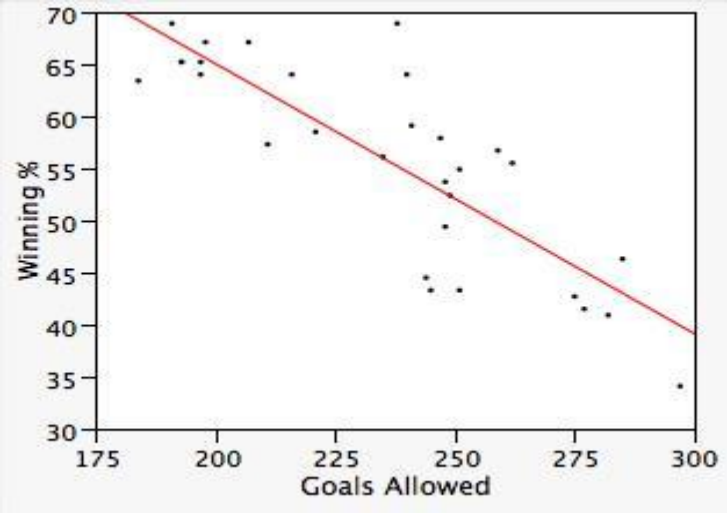
Copyright Macmillan Learning. Powered by Cognero.

Tutorial Ch 2

amount of additive and allowed to set for 24 hours, and then the strength needed to separate the wafers is determined. It is reported that the correlation between strength required and amount of additive was 0.86 pound-force per square inch.
Fill in the blanks: This report is _____ because correlation must be _____.

a. incorrect; unitless
b. correct; positive
c. incorrect; negative
d. None of the above

19. In the National Hockey League, a good predictor of the percent of games won by a team is the number of goals the team allows during the season. Data were gathered for all 30 teams in the NHL, and the scatterplot of their Winning Percent against the number of Goals Allowed in the 2006/2007 season with a fitted least-squares regression line is provided.



The least-squares regression line was calculated to be
Winning Percent (%) = 116.95 – 0.26 * Goals Allowed with $r^2 = 0.69$

Which of the following provides the best interpretation of the slope of the regression line?

a. If the Winning Percent increases by 1%, then the number of Goals Allowed decreases by 0.26.
b. If a team were to allow 100 goals during the season, its Winning Percent would be 90.95%.
c. If Goals Allowed increases by one goal, the Winning Percent increases by 0.26%.
d. If the Winning Percent increases by 1%, then the number of Goals Allowed increases by 0.26%.
e. If Goals Allowed increases by one goal, the Winning Percent decreases by 0.26%.

20. Plots of the residuals versus fits should show a linear pattern if the regression line is a good fit for your data.

a. True
b. False

Copyright Macmillan Learning. Powered by Cognero.

Name:_____ Class:_____ Date:_____

Tutorial Ch 2

Answer Key

1. a

2. c

3. b

4. a

5. a

6. d

7. c

8. b

9. b

10. c

11. a

12. a

13. d

14. b

15. c

16. b

17. c

18. a

19. e

20. b

Copyright Macmillan Learning. Powered by Cognero.